COMPONENT CORRELATIONS IN STRUCTURE-SPECIFIC SEISMIC LOSS ESTIMATION

Brendon A Bradley[*1], Dominic S Lee[2]

[1]Department of Civil and Natural Resources Engineering, University of Canterbury, Private Bag 4800, Christchurch 8020, New Zealand

[2]Department of Mathematics and Statistics, University of Canterbury, Private Bag 4800, Christchurch 8020, New Zealand

*Corresponding author: Ph +64-3-366 7001 ext 7673; Fax: +64-3-364 2758;

Email: bab54@student.canterbury.ac.nz

ABSTRACT

This paper addresses correlations between multiple components in structure-specific seismic loss estimation. To date, the consideration of such correlations has been limited by methodological tractability; increased computational demand; and a paucity of data for their computation. The effect of component correlations, which arise in various forms, is however a significant factor affecting the results of structure-specific seismic loss estimation and therefore it is prudent that adequate consideration is given to their effect. This paper provides details of a tractable and computationally efficient seismic loss estimation methodology in which correlations can be considered. Methods to determine the necessary correlations are discussed, particularly those that can be used in the absence of sufficient empirical data, for which values are suggested based on judgement. The effects of various assumptions regarding correlations are illustrated via application to a case-study office structure. It is observed that certain correlation assumptions can lead to errors in excess of 50% in the lognormal standard deviation in the loss given intensity and loss hazard relationships, while full consideration of partial correlations is 50-times more computationally expensive than other assumptions.

KEYWORDS

Seismic loss estimation; component correlations; generalised equi-correlated model; fragility function correlations; epistemic uncertainty.

INTRODUCTION

Structure-specific seismic loss estimation, in line with the performance-based earthquake engineering (PBEE) framework, involves detailed consideration and computation of losses in a structure due to seismic risk. Uncertainties are explicitly accounted for by treating variables, such as seismic hazard and structural response, probabilistically and integrating over their range of possible values when computing decision variables useful in decision making. Accurate seismic loss estimation requires not only consideration of the uncertainties in the individual components which comprise the structure, but also the correlations between the different components.

Seismic loss estimation methodologies which allow consideration of correlations and the effect of such correlations have received little attention in literature due to, in the author's opinion, several reasons: (i) inevitably their consideration significantly increases the complexity of the algorithms required to perform the loss estimation; (ii) the complicated algorithms significantly increase the computational demand to perform the loss estimation; (iii) there is likely a lack of appreciation for the influence of correlations in the results of the loss estimation.

In the seismic loss estimation framework discussed herein, there are three different correlations which exist between components; these are: (i) correlation between engineering demand parameters (EDP*'*s) for a given intensity measure (IM); (ii) correlation between damage states (DS) given EDP*'s*; and (iii) correlation between component loss (*L*) given DS's. Recent efforts utilizing the Pacific Earthquake Engineering Research (PEER) centre framework, such as Goulet *et al*. [1] and Mitrani-Reiser [2] have considered only expected losses, and therefore neglected correlations (which, for their framework equations, only affects the variance in the loss). Aslani [3], Porter and Kiremidjian [4], Iervolino *et al*. [5], and the ATC-58 guidelines being developed [6] consider correlations explicitly in the $EDP|IM$ relation from the results of seismic response analysis, but Porter and Kiremidjian [4], and ATC-58 [6] do not consider correlations in the $DS|EDP$ and $L|DS$ relationships, while Iervolino *et al*. [5] consider only expected losses and therefore these correlations also do not arise. Aslani [3] and Lee and Kiremidjian [7] consider correlations in the discrete damage state variable using a cumbersome optimisation algorithm, while Baker [8] has recently commented on the use of a more flexible alternative approach. Aslani [3] also considered correlations in the $L|DS$ relationship using correlation data from the construction sector. Aslani [3] and Baker and Cornell [9] both use the first-order second-moment (FOSM) approximation when computing the covariance terms in the total loss because of the perceived computational demand of direct numerical integration. The FOSM approximation has been shown to be of limited accuracy in computing such covariance terms compared to direct evaluation via numerical integration [10]. As far as the author is aware only Aslani [3] has briefly investigated the effect of correlation assumptions on the standard deviation in the total loss given intensity.

The intention of this paper is four-fold. Firstly, the necessary details of a tractable and computationally efficient framework which can account for such correlations are presented. Secondly, significant discussion is given to the determination of the required correlation coefficients, both those which can be computed directly from empirical data, and methods which can be used in the absence of sufficient data. Thirdly, the interaction of correlations and epistemic uncertainties in loss estimation is discussed. Finally, the effects of various assumptions regarding the treatment of correlations are illustrated via an application to a typical office structure.

SEISMIC LOSS ESTIMATION METHODOLOGY CONSIDERING COMPONENT CORRELATIONS

General methodological details

This section presents the mathematical details of a seismic loss estimation method which explicitly accounts for correlations. The basis of the methodology is the PEER performance-based earthquake engineering (PBEE) framework [11] which has been employed by several other researchers [e.g. 1, 3, 9]. Below the various aspects of this framework which are effected by correlations are discussed.

The total loss incurred in a structure when subjected to a ground motion of a specified intensity measure (IM) is conditioned on the mutually exclusive and collectively exhaustive events of collapse and non-collapse. From the total expectation theorem [12], the mean, $\mu_{L_T|IM}$, and variance, $\sigma^2_{L_T|IM}$, of the total loss given IM are given by [e.g. 3]:

$$\mu_{L_T|IM}(im) = \mu_{L_T|IM,NC}(im)\left(1 - P_{C|IM}(im)\right) + \mu_{L_T|C} P_{C|IM}(im) \tag{1}$$

$$\begin{aligned}\sigma^2_{L_T|IM}(im) = &\ \sigma^2_{L_T|IM,NC}(im)\left(1 - P_{C|IM}(im)\right) + \sigma^2_{L_T|C} P_{C|IM}(im) \\ &+ \left(\mu_{L_T|IM,NC}(im) - \mu_{L_T|C}\right)^2 P_{C|IM}(im)\left(1 - P_{C|IM}(im)\right)\end{aligned} \tag{2}$$

where $\mu_{L_T|IM,NC}(im)$, $\sigma^2_{L_T|IM,NC}(im)$, $\mu_{L_T|C}$, and $\sigma^2_{L_T|C}$ are the mean and variance in the loss given no collapse and collapse, respectively; and $P_{C|IM}(im)$ is the probability of collapse given $IM = im$. Given collapse occurs it is assumed that the mean and variance in the total loss are independent of IM (i.e. $\mu_{L_T|IM,C} = \mu_{L_T|C}$).

In the case of no collapse, the total loss is comprised of the sum of the loss to individual components at spatially different locations throughout the structure, with mean, $\mu_{L_T|IM,NC}$, and variance, $\sigma^2_{L_T|IM,NC}$, given by:

$$\mu_{L_T|IM,NC}(im) = \sum_{i=1}^{N_C} \mu_{L_i|IM,NC}(im) \tag{3}$$

$$\sigma^2_{L_T|IM,NC}(im) = \sum_{i=1}^{N_C} \sigma^2_{L_i|IM,NC}(im) + 2\sum_{i=1}^{N_C}\sum_{j=1}^{i-1} \sigma_{L_i,L_j|IM,NC}(im) \tag{4}$$

where $\mu_{L_i|IM,NC}(im)$ and $\sigma^2_{L_i|IM,NC}(im)$ are the mean and variance, respectively, in the loss to component i given $IM = im$; $\sigma_{L_i,L_j|IM,NC}(im)$ is the covariance in the loss between components i and j given $IM = im$; and N_C is the number of different components in the structure.

In the case of collapse, the total loss is given as the sum of the cost to replace all of the components in the structure (whether they are damaged or not), and also additional costs to account for re-design and demolition [3], making it potentially significantly different from its current market value [13]. The mean, $\mu_{L_T|C}$, and variance, $\sigma^2_{L_T|C}$ in the total loss given collapse are given by:

$$\mu_{L_T|C} = (1 + C_{RDD})\sum_{i=1}^{N_{CCI}} \mu_{CCI_i} \tag{5}$$

$$\sigma^2_{L_T|C} = (1 + C_{RDD})^2\left[\sum_{i=1}^{N_{CCI}} \sigma^2_{CCI_i} + 2\sum_{i=1}^{N_{CCI}}\sum_{j=1}^{i-1} \rho_{CCI_i,CCI_j}\sigma_{CCI_i}\sigma_{CCI_j}\right] \tag{6}$$

where μ_{CCI_i} and $\sigma^2_{CCI_i}$ are the mean and variance in the cost of construction cost item (CCI) i; N_{CCI} is the number of construction cost items involved in the construction of the structure; ρ_{CCI_i,CCI_j} is the correlation between the cost of construction cost items i and j; and C_{RDD} is additional costs due to redesign and demolition as a proportion of the total cost. Note the mean and variance in the loss to repair or replace each component should obviously account for contractor overhead, inflation and location [14].

Equations (1)-(6) reveal that given the above assumptions, correlations between component losses (which appear in Equations (4) and (6)) only affect the variance in the total loss, and therefore do not need to be considered when computing only the expected value of the loss [e.g. 1, 2]. As will be shown however, the magnitude of the variance in the total loss is such that, in the author's opinion, it should always be considered when making earthquake risk decisions.

To rationally determine the correlation coefficient in the total loss given no collapse, it is necessary to further examine the covariance in the loss between components i and j, $\sigma_{L_i,L_j|IM,NC}$. Using the general relationship between covariance and expectations [e.g. 12], $\sigma_{L_i,L_j|IM,NC}$ can be expressed as (where the conditioning on no collapse, NC, has been dropped where obvious for brevity):

$$\sigma_{L_i,L_j|IM,NC}(im) = \mu_{L_iL_j|IM}(im) - \mu_{L_i|IM}(im)\mu_{L_j|IM}(im) \tag{7}$$

where $\mu_{L_iL_j|IM}(im)$ is the expected value of the product L_iL_j given $IM = im$; and $\mu_{L_i|IM}(im)$ and $\mu_{L_j|IM}(im)$ are the expected losses of components i and j given $IM = im$, respectively. $\mu_{L_i|IM}$, and similarly $\mu_{L_j|IM}$, are computed by:

$$\mu_{L_i|IM}(im) = \int \mu_{L_i|EDP_i}(edp_i) f_{EDP_i|IM}(edp_i|im) dEDP_i \tag{8}$$

where EDP_i is the engineering demand parameter that component i is subjected to; $\mu_{L_i|EDP_i}(edp)$ is the expected loss to component i given $EDP_i = edp_i$; and $f_{EDP_i|IM}(edp_i|im)$ is the probability density function (pdf) of EDP_i given $IM = im$. Equation (8) is an application of the total probability theorem [e.g. 12] and makes the conditional independence assumption that conditioned on EDP_i the mean (and generally the distribution) of L_i is independent of IM.

Using the same assumptions in Equation (8), the first term on the right-hand side of Equation (7), $\mu_{L_iL_j|IM}$, is given by:

$$\mu_{L_iL_j|IM}(im) = \iint \mu_{L_iL_j|EDP_i,EDP_j}(edp_i, edp_j) f_{EDP_i,EDP_j|IM}(edp_i, edp_j|im) dEDP_i dEDP_j \tag{9}$$

where $\mu_{L_iL_j|EDP_i,EDP_j}(edp_i, edp_j)$ is the expected value of the product L_iL_j given $EDP_i = edp_i$ and $EDP_j = edp_j$; and $f_{EDP_i,EDP_j|IM}(edp_i, edp_j|im)$ is the bi-variate pdf of EDP_i and EDP_j given $IM = im$.

The expected loss in component i given $EDP_i = edp_i$, $\mu_{L_i|EDP_i}$ can be determined via the use of discrete damage states (DS's) as:

$$\mu_{L_i|EDP_i}(edp_i) = \sum_{k=1}^{N_{DS,i}} \mu_{L_i|DS_k}(ds_k) P_{DS_k|EDP_i}(ds_k|edp_i) \tag{10}$$

where $\mu_{L_i|DS_k}(ds_k)$ is the expected loss in component i given $DS_k = ds_k$; $P_{DS_k|EDP_i}(ds_k|edp_i)$ is the probability of $DS_k = ds_k$ given $EDP_i = edp_i$; and $N_{DS,i}$ is the number of damage states for component i. Again, in Equation (10) the total probability theorem is used as well as the conditional independence assumption (i.e. that given DS_k, L_i is independent of EDP_i). Details on loss and fragility functions which are needed to determine $\mu_{L_i|DS_k}$ and $P_{DS_k|EDP_i}$ can be found in, for example, Mitrani-Reiser [2] and Porter *et al.* [15].

Similar to Equation (10), $\mu_{L_iL_j|EDP_i,EDP_j}$ can be computed by:

$$\begin{aligned}\mu_{L_iL_j|EDP_i,EDP_j}(edp_i, edp_j) &= \sum_{k=1}^{N_{DS,i}}\sum_{l=1}^{N_{DS,j}} \mu_{L_iL_j|DS_k,DS_l}(ds_k, ds_l) P_{DS_k,DS_l|EDP_i,EDP_j}(ds_k, ds_l | edp_i, edp_j) \\ &= \sum_{k=1}^{N_{DS,i}}\sum_{l=1}^{N_{DS,j}} \left(\mu_{L_i|DS_k}(ds_k)\mu_{L_j|DS_l}(ds_l) + \sigma_{L_i,L_j|DS_k,DS_l}(ds_k, ds_l)\right) x \\ &\qquad P_{DS_k,DS_l|EDP_i,EDP_j}(ds_k, ds_l | edp_i, edp_j)\end{aligned} \tag{11}$$

where $P_{DS_k,DS_l|EDP_i,EDP_j}(ds_k, ds_l | edp_i, edp_j)$ is the (joint) probability of $DS_k = ds_k$ and $DS_l = ds_l$ in components i and j given $EDP_i = edp_i$ and $EDP_j = edp_j$, respectively; and $\sigma_{L_i,L_j|DS_k,DS_l}(ds_k, ds_l)$ is the covariance in the loss in components i and j given $DS_k = ds_k$ and $DS_l = ds_l$. The term in parentheses on the second line of Equation (11) is obtained from the first using an equivalent form of Equation (7).

Equations (1)-(11) completely define those aspects of the seismic loss estimation methodology used here which involve correlations. Other equations which comprise the methodology not involving correlations can be found in Bradley *et al*. [16]. From Equations (1)-(11), it can be observed that the effects of correlations affect four different terms, namely: $f_{EDP_i,EDP_j|IM}$ (Equation (9)), $P_{DS_k,DS_l|EDP_i,EDP_j}$ (Equation (11)), $\sigma_{L_i,L_j|DS_k,DS_l}$ (Equation (11)), and ρ_{CCI_i,CCI_j} (Equation (6)). These four terms are dependent on the correlations in the *EDP*|*IM*, *DS*|*EDP* and *L*|*DS* relationships.

Note that under the adopted framework only a scalar ground motion intensity measure, *IM*, is considered. Should a vector intensity measure be considered [17], then correlations between the individual *IM* terms should also be accounted for.

Correlations in the *EDP*|*IM* relationship

The previous section illustrated that the effect of correlations in the *EDP*|*IM* relationship appears in the joint-distribution, $f_{EDP_i,EDP_j|IM}$. As the marginal ln*EDP*|*IM* distribution, $f_{\ln EDP_i|IM}$, is typically assumed to have a normal distribution [18], (i.e. $f_{EDP_i|IM}$ has a lognormal distribution) then it is (reasonably) assumed that the joint distribution, $f_{\ln EDP_i,\ln EDP_j|IM}$, is well represented by a bi-variate normal distribution (i.e. $f_{EDP_i,EDP_j|IM}$ has a bi-variate lognormal distribution). Thus $f_{\ln EDP_i,\ln EDP_j|IM}$ is given by:

$$\begin{aligned}f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{X,Y}^2}}\exp\Bigg[\frac{-1}{2\left(1-\rho_{X,Y}^2\right)}\Bigg\{\left(\frac{x-\mu_X}{\sigma_X}\right)^2 \\ -2\rho_{X,Y}\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\Bigg\}\Bigg]\end{aligned} \tag{12}$$

where $X = \ln EDP_i|IM$; $Y = \ln EDP_j|IM$; $\mu_X = \mu_{\ln EDP_i|IM}$ and $\sigma_X = \sigma_{\ln EDP_i|IM}$ are the mean and standard deviation of the ln$EDP_i|IM$ relation; and $\rho_{X,Y} = \rho_{\ln EDP_i|IM,\ln EDP_j|IM}$ is the correlation coefficient between ln$EDP_i|IM$ and ln$EDP_j|IM$.

Correlations in *DS|EDP* relationship

Correlations in the *DS|EDP* relationship appear in the joint probability $P_{DS_k,DS_l|EDP_i,EDP_j}$. Determination of this joint probability is complicated by the fact that *DS* is a discrete variable. Aslani [3] and Lee and Kiremidjian [7] propose the use of an optimisation procedure to determine $P_{DS_k,DS_l|EDP_i,EDP_j}$ for a given correlation. Baker [8] notes that this optimisation approach is cumbersome and instead proposes that fragility functions (which define the *DS|EDP* relationship) be considered in terms of the (continuous) EDP which causes the (discrete) DS rather than the probability of DS given EDP. Baker [8] then illustrates how Monte Carlo simulation can be used to determine $P_{DS_k,DS_l|EDP_i,EDP_j}$. Here the approach of considering the fragility function as a continuous function of capacity is taken, but solution via a tractable analytical approach is developed which is significantly more computationally efficient than simulation-based methods.

Firstly, it is noted that for a single component the damage state *k* probability, $P_{DS_k|EDP_i}$, is given by the difference in the fragility functions defining the probability of exceeding DS_k and DS_{k+1} [e.g. 19]:

$$\begin{aligned} P_{DS_k|EDP_i}\left(ds_k\left|edp_i\right.\right) &= F_{DS_k|EDP_i}\left(ds_k\left|edp_i\right.\right) - F_{DS_{k+1}|EDP_i}\left(ds_{k+1}\left|edp_i\right.\right) \\ &= F_{DS_k|EDP_i}\left(ds_k\left|edp_i\right.\right) - \sum_{m=k+1}^{N_{DS,i}} P_{DS_m|EDP_i}\left(ds_m\left|edp_i\right.\right) \end{aligned} \tag{13}$$

where the fragility function, $F_{DS_k|EDP_i}$, is typically assumed to have a lognormal distribution [15]. Use of Equation (13) in the case of $k = N_{DS}$ can be handled simply by defining $F_{DS_{N_{DS+1}}|EDP_i} = 0$. The first line of Equation (13) is the conventional form by which the damage state probability is defined [e.g. 19], while the equivalent second line is introduced here to aid in the description of the following paragraph. The second line of Equation (13) describes the probability of DS_k given EDP_i literally as the probability of DS_k or greater (i.e. $F_{DS_k|EDP_i}$), less all the greater terms (those in the summation).

Using the same logic for the single variable case as in the second line of Equation (13), it is trivial to show that the joint probability of DS_k and DS_l given EDP_i and EDP_j is given by:

$$\begin{aligned} P_{DS_k,DS_l|EDP_i,EDP_j}\left(ds_k, ds_l\left|edp_i, edp_j\right.\right) &= F_{DS_k,DS_l|EDP_i,EDP_j}\left(ds_k, ds_l\left|edp_i, edp_j\right.\right) \\ &\quad - \sum_{m=k}^{N_{DS,i}} \sum_{\substack{n=l \\ n\neq l \text{ if } m=k}}^{N_{DS,j}} P_{DS_m,DS_n|EDP_i,EDP_j}\left(ds_m, ds_n\left|edp_i, edp_j\right.\right) \end{aligned} \tag{14}$$

Figure 1 illustrates schematically the implications of Equation (14) for a particular case in which $k = 2$, $l = 2$, $N_{DS,i} = 3$, $N_{DS,j} = 4$. Note that in order to compute Equation (14), one must have computed all terms in the summation *a priori*. The numbered square brackets in Figure 1 illustrate one possible sequence by which $P_{DS_{k=2},DS_{l=2}|EDP_i,EDP_j}$ can be determined. For example, in step [1], Equation (14) becomes: $P_{DS_{k=3},DS_{l=4}|EDP_i,EDP_j} = F_{DS_{k=3},DS_{l=4}|EDP_i,EDP_j}$ (i.e. all the terms in the summation are zero). In step [2], Equation (14) becomes: $P_{DS_{k=2},DS_{l=4}|EDP_i,EDP_j} = F_{DS_{k=2},DS_{l=4}|EDP_i,EDP_j} - P_{DS_{k=3},DS_{l=4}|EDP_i,EDP_j}$, where the second term was previously evaluated in step [1]. Thus at each step only $F_{DS_k,DS_l|EDP_i,EDP_j}$ must be computed. It is also important to note that for each component pair (i.e. each *i,j* pair), Equation (11) requires $P_{DS_k,DS_l|EDP_i,EDP_j}$ for all $k = 1\text{-}N_{DS.i}$ and $l = 1\text{-}N_{DS.j}$; thus the sequential process in Equation (14) does not result in any unnecessary computations.

In the case of a single component, the damage state fragility function, $F_{DS_k|EDP_i}(ds_k|edp_i) = P(DS_k \geq ds_k|EDP_i = edp_i)$, which can be interpreted as the probability of $DS_k \geq ds_k$ given $EDP_i = edp_i$ is equivalent to $P(C_{i,k} < EDP_i)$, the probability that the demand, EDP_i is greater than the damage state k capacity of component i, $C_{i,k}$ [8]. As $F_{DS_k|EDP_i}(ds_k|edp_i)$ is typically defined by a lognormal distribution [15] then it follows that $P(C_{i,k} < EDP_i)$ also has a lognormal distribution, and $P(\ln C_{i,k} < \ln EDP_i)$ a normal distribution. If it is (reasonably) assumed that $F_{DS_k,DS_l}(\ln edp_i, \ln edp_j)$ is a cumulative bi-variate normal distribution then it follows by definition that:

$$
\begin{aligned}
&F_{DS_k,DS_l|\ln EDP_i,\ln EDP_j}(ds_k, ds_l|\ln edp_i, \ln edp_j) \\
&\quad = P(DS_k \geq ds_k, DS_l \geq ds_l|\ln EDP_i = \ln edp_i, \ln EDP_j = \ln edp_j) \\
&\quad = P(\ln C_{i,k} < \ln edp_i, \ln C_{j,l} < \ln edp_j) \\
&\quad = \iint\limits_{\substack{\ln c_{i,k} < \ln EDP_i \\ \ln c_{j,l} < \ln EDP_j}} f_{\ln C_{i,k},\ln C_{j,l}}(\ln c_{i,k}, \ln c_{j,l})\, d\ln C_{i,k}\, d\ln C_{j,l}
\end{aligned}
\tag{15}
$$

where $f_{\ln C_{i,k},\ln C_{j,l}}$ is a bi-variate normal distribution pdf (i.e. the same form as that given in Equation (12)) of the component capacities. Thus, only the correlation coefficient, $\rho_{\ln C_{i,k},\ln C_{j,l}}$, defining the correlation between the damage state k and l capacities of components i and j, respectively, is required in addition to the conventional fragility function data (which defines the marginal mean and variances).

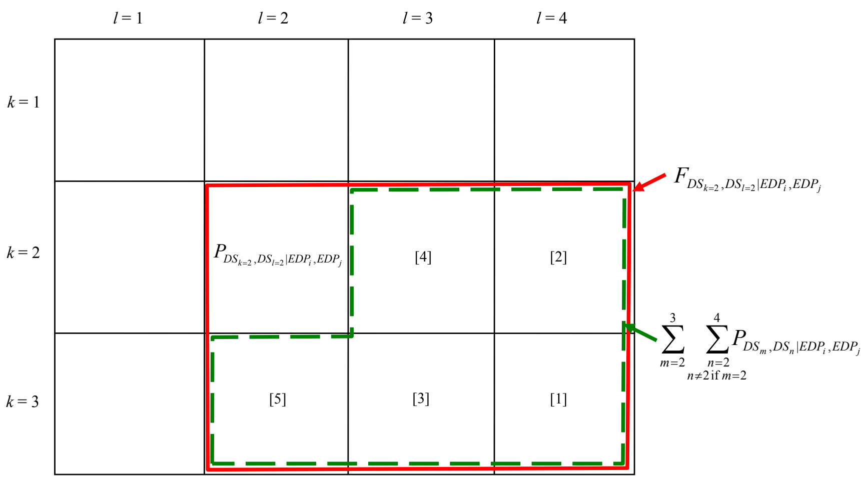

Figure 1: Schematic illustration of the computation of damage state probabilities as defined in Equation (14) for the case of $k = 2$, $l = 2$, $N_{DS,i} = 3$, $N_{DS,j} = 4$.

Equation (15) at first may appear to be computationally demanding because of the double integral. However, because of the frequent use of the cumulative bi-variate normal distribution in probability theory, highly efficient numerical algorithms are available. Such algorithms transform the double integral into a single integral and obtain the solution using a combination of analytical integration and as little as 4 Gauss quadrature points [20, 21]. Thus, it can be appreciated that determination of $P_{DS_k,DS_l|EDP_i,EDP_j}$ via Equations (14) and (15) is orders of magnitude more efficient than the optimization algorithm of Lee and Kiremidjian [7], or solution by Monte Carlo simulation as discussed by Baker [8].

An additional benefit of the formulation given by Equations (14) and (15) over the optimization approach discussed in Lee and Kiremidjian [7], is that the damage state correlation, $\rho_{\ln C_{i,k},\ln C_{j,l}}$, can potentially be a function of both component types and damage state numbers, while the optimization algorithm allows only a single correlation coefficient per component pair.

Correlations in the *L|DS* relationship

Correlations between the cost to repair damage to different components appear in the term $\sigma_{L_i,L_j|DS_k,DS_l} = \rho_{L_i,L_j|DS_k,DS_l}\sigma_{L_i|DS_k}\sigma_{L_j|DS_l}$ given in Equation (11), where $\rho_{L_i,L_j|DS_k,DS_l}$ is the correlation in the loss to components *i* and *j*, due to damage states *k* and *l*, respectively. Correlations in the cost to replace (equivalent to 'repairing' a component in its failure damage state) individual components when collapse occurs is given by ρ_{CCI_i,CCI_j} in Equation (6).

Neglected correlations

The previous sections address correlations which appear in the structure-specific seismic loss estimation framework given by Equations (1)-(10). There are however additional, potentially important, correlations which are not considered because of the assumptions made in the framework. In particular, the conditional independence assumption, which allows the *L|DS*, *DS|EDP*, *EDP|IM*, and λ_{IM} relationships to be treated independently, and combined using the total probability theorem, means that correlations are only considered within, and not between, these relationships. For example, uncertainty in the capacity of a component of the lateral load resisting system of a structure affects both the damageability of the component itself, but also the dynamic response of the entire structural system. While this uncertainty can be separately considered for the *EDP|IM* and *DS|EDP* relationships (using a stochastic seismic response model and a component fragility curve, respectively), since they are characterised separately, the correlation between the *EDP|IM* and *DS|EDP* relationships due to the random capacity for this single component is not correctly accounted for in the strict sense. This is one acknowledged drawback of the conditional independence assumption.

CAUSES OF AND METHODS TO DETERMINE CORRELATIONS

Correlation in the *EDP|IM* relationship

Correlations between different EDP's for a given IM (i.e. $\rho_{\ln EDP_i|IM,\ln EDP_j|IM}$) occur due to the dynamic characteristics of the structure and ground motion which it is subjected to. This correlation structure is indeed complex, but can however be determined from the results of multiple time-history or modal pushover [e.g. 22] analyses which are ground-motion dependent. Given a suite of N_{gm} ground motions scaled to a specific value of *IM*, seismic response analyses will yield an N_{edp} x N_{gm} matrix of seismic response, ***EDP*** (i.e. one column for each ground motion and one row for each EDP_i value being monitored), the correlation coefficient for EDP_i and EDP_j can then be computed by:

$$\rho_{\ln EDP_i|IM,\ln EDP_j|IM} = \sum_{k=1}^{N_{gm}} \left(\frac{\ln EDP_{i,k} - \mu_{\ln EDP_i}}{\sigma_{\ln EDP_i}} \right) \left(\frac{\ln EDP_{j,k} - \mu_{\ln EDP_j}}{\sigma_{\ln EDPj}} \right) = \sum_{k=1}^{N_{gm}} \varepsilon_{i,k} \varepsilon_{j,k} \tag{16}$$

where $EDP_{i,k} = \boldsymbol{EDP}(i,k)$ is the value of i^{th} EDP monitored due to ground motion k; $\mu_{\ln EDP_i}$ and $\sigma_{\ln EDP_i}$ are the mean and standard deviation of $\ln EDP_i$ over the N_{gm} different ground motions; and $\varepsilon_{i,k}$ is the so-called standardized residual of $\ln EDP_i$.

The correlations of the N_{edp} different EDP values being monitored are defined by a $N_{edp} \text{ x } N_{edp}$ symmetric correlation matrix with $0.5\, N_{edp}\, (N_{edp} - 1)$ unique correlation coefficients. Figure 2 illustrates the lower-triangular portion of the correlation matrix based on the seismic response analyses of the 10 storey office building discussed in Bradley *et al.* [23, 24]. In Figure 2, *EDP* numbers 1-10 are the peak interstorey drift ratios on floors 1-10, and *EDP* numbers 11-21 are peak floor accelerations on the 1st – roof floors. Although the correlation structure is clearly complex, three features can be observed. Firstly, the lower triangular portion of the correlation matrix can be distinguished into three sections: correlations between two peak interstorey drift EDPs (i.e. EDP_i and $EDP_j \leq 10$); correlations between two peak floor accelerations (i.e. EDP_i and $EDP_j \geq 11$) and correlations between a peak interstorey drift and peak floor acceleration (i.e. $EDP_i \leq 10$ and $EDP_j \geq 11$). The magnitude of the correlations in each of these three sections can be put in descending order as: peak floor acceleration vs. peak floor acceleration; peak interstorey drift vs. peak interstorey drift; peak interstorey drift vs. peak floor acceleration. Finally, within any of these three sections, as the value of $\left|EDP_i - EDP_j\right|$ increases, there is a trend for the correlation to reduce.

While the above three observations are insightful, if the correlation matrix can be obtained directly from the results of time-history analyses then such observations are only useful for validation with intuition. Various simplified methods have however been proposed to determine the *EDP*|*IM* relationship for a structure which are ground motion independent [25]. Such ground-motion independent methods therefore do not enable the computation of *EDP*|*IM* correlations as given in Equation (16). In the following paragraphs a simple *EDP*|*IM* correlation model is developed for multi-storey buildings based on the structural analysis results of Bradley *et al.* [23, 24] which can be used in conjunction with such ground motion independent simplified methods.

Figure 3a and 3b illustrate the relationship between the correlation coefficient and the number of floors separation, n_{fs}, (equivalent to $\left|EDP_i - EDP_j\right|$) for peak interstorey drifts and peak floor accelerations, respectively. In Figure 3a and 3b, each point is one correlation coefficient, and each dashed line is the arithmetic mean of the correlations for a specific IM level. The 50 ground motions, 21 EDPs and 9 IM levels used in Bradley *et al.* [23, 24] gives a total of 405 (i.e. 0.5x10x(10 - 1)x9) and 495 correlation coefficients for the peak interstorey drifts and peak floor accelerations, respectively. The similar trends observed in the 9 different arithmetic means in each plot suggests that their is not an overly significant variation in the correlation vs. n_{fs} trend for ground motion intensities resulting in elastic response through to collapse [23, 24]. The solid lines provide simple piecewise linear fits to the data with equation inset in each figure. Figure 3c illustrates the correlations coefficient values between a peak interstorey drift and peak floor acceleration, ρ_{a_i,θ_j}, as a function of n_{fs}. It can be seen that the scatter for ρ_{a_i,θ_j} is notably larger than that for ρ_{θ_i,θ_j} and ρ_{a_i,a_j} in Figures 3a and 3b, suggesting that the correlation coefficient, ρ_{a_i,θ_j}, is significantly dependent other variables in addition to n_{fs}. Despite the larger scatter in Figure 3c a simplified value of $\rho_{a_i,\theta_i} = 0.4$ is suggested.

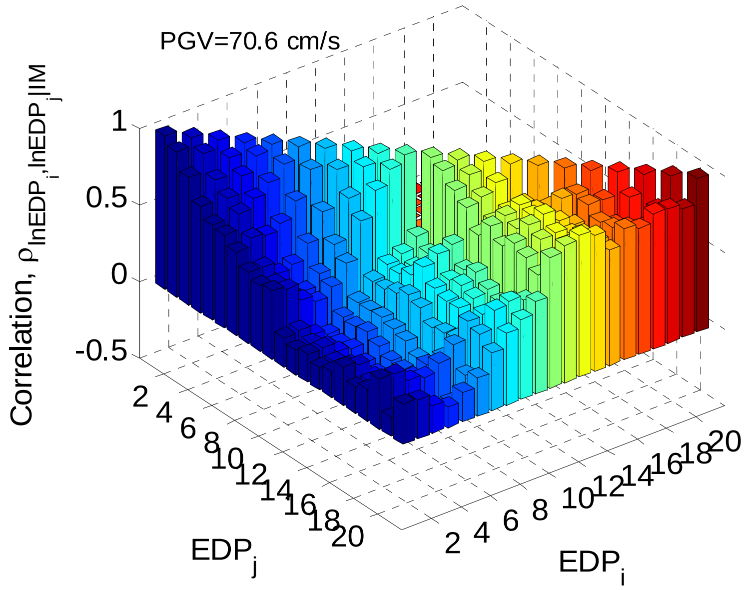

Figure 2: Typical correlation matrix of the *EDP|IM* relationship for the case study structure. EDP_1-EDP_{10} are peak interstorey drift ratios on floors 1-10, and EDP_{11}-EDP_{21} are peak floor accelerations on the 1st – roof floors.

Clearly the above simplified relationships are tentative in the sense that only the seismic response analyses of a single capacity-designed structure have been used in their development. That said, Baker and Cornell [26] also illustrate the drift vs. drift correlation as a function of floor separation for a seven storey structure, and the correlation vs. floor separation trend is remarkably similar to that presented here. This highly simplified method which provides a partial correlation value with no detailed information required is likely to be at a consistent level of accuracy compared to the simplified seismic response analysis used, and is therefore a plausible model until further studies are conducted.

Correlation in the *DS|EDP* relationship

Correlation in the *DS|EDP* relationships of different components occurs due to the shared uncertainties in the EDP values at which the specific DS's occur. This uncertainty can be separated into shared uncertainty due to the EDP definition and the component capacity.

While the conditional independence assumption is made in the loss assessment framework for simplicity and tractability (i.e. that given *EDP* the damage state probability is independent of *IM*), common EDP's, such as peak interstorey drift and peak floor acceleration do not account for the frequency content and duration of the seismic excitation. Thus, for example, a large magnitude earthquake (with a correspondingly long duration) is likely to lead to more damage throughout the structure than a small magnitude earthquake (with short duration) which causes the same (peak) demand vector, ***EDP*** = ***edp***, but a different response history [27] (i.e. the EDP used is an *insufficient* [28] predictor of DS).

Shared uncertainty in the capacity of the different components is present due to the dependence in the material properties which the components are comprised of, and the similarity in the installation techniques required.

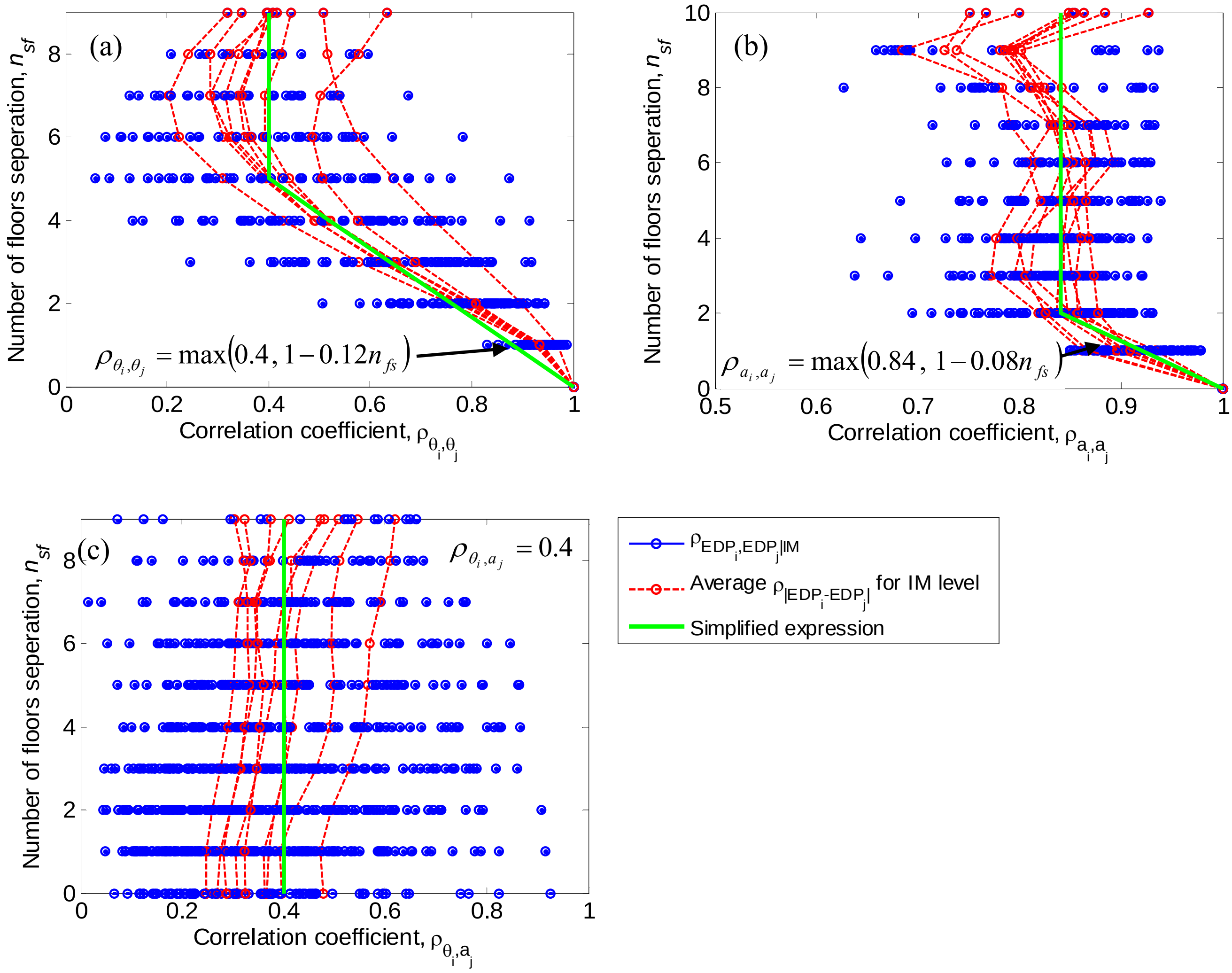

Figure 3: Trends in the EDP|IM correlations as a function of the number of floors separation of the demand being measured: (a) correlations between peak interstorey drifts on different floors; (b) correlations between peak floor accelerations on different floors; and (c) correlations between peak interstorey drift and peak floor acceleration on the different floors.

Unlike the correlations in the *EDP*|*IM* relationship, which, for a specific structure, can be obtained from the results of ground-motion dependent numerical simulations, there is a paucity of data for use in determining correlations between *DS*|*EDP* relations. Lee and Kiremidjian [7] and Aslani [3] both discuss the consideration of *DS*|*EDP* correlations but do not present any means or data to determine the correlations. Baker and Cornell [9] propose the use of the generalised equi-correlated model [e.g. 29] for determining the correlation in a 'collapsed' *L*|*EDP* relationship. Here the generalised equi-correlated model is developed for the *DS*|*EDP* relationship (and also the *L*|*DS* relationship in the next section). Baker [8] notes that there are several methods by which correlations can be (in future) determined from observations. It is suggested therefore that a generalised equi-correlated model is used to initially determine correlations, and as observational data becomes available in the future they can be used in a Bayesian framework to update specific correlation values.

The generalised equi-correlated model adopted for the *DS*|*EDP* relationship is composed of five mutually independent variables. These five variables have been chosen to account for the main contributing sources of uncertainty, and can be easily handled in the loss assessment computation (i.e. a more robust model would further break down some of these variables, but the input and computational requirements to compute the correlation would be significantly increased). The five variables considered represent: (i) uncertainty due to the EDP definition common to all components (e.g. the aforementioned example relating to duration of shaking); (ii) uncertainty due to the EDP definition common to components subjected to the same EDP (e.g. two components both sensitive to 2nd floor acceleration); (iii) uncertainty in seismic capacity common to components made of the same material; (e.g. a structural beam and structural wall both made of concrete); (iv) uncertainty in seismic capacity common to components of the same type (e.g. two different concrete structural beams); and (v) uncertainty in seismic capacity of a single component independent of all other components).

Based on the five variables explained above, the total uncertainty in the *DS*|*EDP* relationship for component *i*, which is dependent on EDP_k is given by:

$$\sigma^2_{\ln EDP_{k_i}|DS} = \left(\sigma^2_{structure_i} + \sigma^2_{EDP_{k_i}}\right) + \left(\sigma^2_{mat_i} + \sigma^2_{comptype_i} + \sigma^2_{comp_i}\right) \tag{17}$$

where $\sigma^2_{structure_i}$ and $\sigma^2_{EDP_{k_i}}$ are the shared uncertainties due to the EDP definition for the entire structure, and EDP_k, respectively for component *i* (due to insufficiency of the EDP's used); $\sigma^2_{mat_i}$, $\sigma^2_{comptype_i}$, and $\sigma^2_{comp_i}$ are the shared uncertainties in the capacity for material type, component type, and component, respectively, for component *i*. It is assumed that each of the five variables is either independent or perfectly correlated with the same variable of a different component.

It should be noted that some fragility functions have been developed without consideration of the insufficiency of the EDP, typically those based on quasi-static laboratory experiments [30]. In such cases, the total uncertainty represents only capacity uncertainty. Further research is needed to more accurately determine the magnitude of the uncertainty due to using insufficient EDP's. Alternatively, using EDP's which account for intensity, frequency content and duration of excitation can be expected to significantly reduce the demand uncertainty portion of the fragility function uncertainty.

A limitation of the equi-correlated model as adopted in this application is that no consideration has been made of the particular damage states of the components. For example, the dependence between cracking damage states in a concrete wall and concrete beam (both dependent on the concrete tensile strength) will have a higher correlation that cracking in the concrete wall and failure in the concrete beam (failure in the ductile beam being primarily dependent on the reinforcing steel properties). As no empirical data is currently available to warrant a damage state dependent correlation such an effort is not pursued here, although it is again noted that mathematically speaking a damage state dependent correlation is not a problem for the framework presented in this manuscript.

From Equation (17), and using the relation, $Cov[a,b+c] = Cov[a,b] + Cov[a,c]$, the covariance between two different components, *i* and *j*, is given by:

$$\begin{aligned} Cov\left[\ln EDP_{k_i}|DS, \ln EDP_{k_j}|DS\right] = &\left(\sigma_{structure_i}\sigma_{structure_j} + \delta_{k_i k_j}\sigma_{EDP_{k_i}}\sigma_{EDP_{k_j}}\right) \\ &+ \left(\delta_{mat_i mat_j}\sigma_{mat_i}\sigma_{mat_j} + \delta_{comptype_i comptype_j}\sigma_{comptype_i}\sigma_{comptype_j} + \delta_{ij}\sigma_{comp_i}\sigma_{comp_j}\right) \end{aligned} \tag{18}$$

where δ_{ij} is the Kronecker delta function equal to one if $i = j$ and zero otherwise (i.e. corresponding to zero of perfect correlations). Note that no Kronecker delta function is necessary for the product $\sigma_{structure_i}\sigma_{structure_j}$, since it is common to all components and thus is always equal one.

Based on the author's judgement, of the total variance in the damage state uncertainty (Equation (17)): 30% is assumed to be due to seismic demand uncertainty and 70% due to component capacity uncertainty; 67% of the demand uncertainty is assumed to be common to the entire structure ($\sigma^2_{structure_i}$), and 33% to a specific EDP ($\sigma^2_{EDP_k}$); 50% of the capacity uncertainty is assumed to be common to specific material types ($\sigma^2_{mat_i}$), 35% common to specific component types ($\sigma^2_{comptype_i}$) and 15% specific to each component ($\sigma^2_{comp_i}$). If, in Equations (18), each of the five variables are written as a proportion of the total uncertainty (e.g. $\sigma^2_{structure_i} = 0.3\left(0.67\sigma^2_{\ln EDP_{k_i}|DS}\right)$) then the correlation coefficient, becomes, after simplification:

$$\rho_{\ln EDP_{k_i}|DS,\ln EDP_{k_j}|DS} = 0.3\left(0.67 + 0.33\delta_{k_i k_j}\right) + 0.7\left(0.5\delta_{mat_i mat_j} + 0.35\delta_{comptype_i comptype_j} + 0.15\delta_{ij}\right) \quad (19)$$

As illustrated by Equation (19), the assumption that the standard deviation values for each of the five variables are a function of the total uncertainty in the fragility function, as opposed to $\sigma^2_{structure_i}$ being the same for all components, means that the correlation is in fact dependent only on the four δ_{ij} terms. The argument against use of constant values for the variables is based on the likelihood that different components will be less or more sensitive to different variables. The fact that some fragility functions have (logarithmic) standard deviations as low as 0.28 considering both demand and capacity portions of the uncertainty [e.g. 31], while others can be in the vicinity of 0.6 without considering the demand portion of the uncertainty [e.g. 19] suggests that the magnitude of the variable uncertainties is not likely to be the same for different components.

Correlation in the *L|DS* relationship

Correlations between the *L|DS* relationships of different components arise due to similarity in the repair actions required to repair the component. This similarity will dictate, for example, whether the same labourers repair both components *i* and *j*, and whether the same materials are required to conduct the repair. Both of these effects will directly influence the correlation in the loss (be it direct repair cost or time taken to perform the repairs).

While the consideration of such correlations for seismic loss estimation is a relatively new problem, correlations between construction costs for determining total project budgets and cost contingencies have received attention in engineering management literature [e.g. 32, 33]. Not surprisingly, the same problem of limited empirical data to determine such correlations is the central issue. Touran and Wiser [34] provide correlation coefficients between unit costs of general construction items such as concrete, metals, electrical, etc. based on data from 26 projects, which Aslani [3] used in seismic loss estimation. Apart from a correlation of 0.79 between electrical and mechanical costs, all correlation values in Touran and Wiser [34] are below 0.51, most likely due to the broad construction category definitions used. An alternative approach adopted by Hudak and Maxwell [33] is to use a so-called macro approach in which common external (or macro-) factors are identified which are common to multiple cost items. The macro approach of Hudak and Maxwell [33] is similar to the aforementioned generalised equi-correlated model used for the *DS|EDP* correlations, but allows the use of coefficients for each of the macro factors, and also partial correlation between the same macro-factors in different components. In comparison, coefficients of one, and either none or perfect correlations between the variables were used in the generalised equi-correlated model used for the *DS|EDP* correlations. While the model of Hudak and Maxwell [33] is more general than the model used herein, it also requires additional data (i.e. values of the coefficients and partial correlations). Thus, such a model may be more appropriate in future, when additional data warrant such generalisation.

The generalised equi-correlated model adopted for the L|DS relationship is composed of three variables, which represent: (i) uncertainty in repair cost/duration common to components made of the same material; (e.g. a structural beam and structural wall both made of concrete); (ii) uncertainty in repair cost/duration common to components of the same type (e.g. two different concrete structural beams); and (iii) uncertainty in repair cost/duration of a single component independent of all other components). Based on these three variables, the total uncertainty in the L|DS relationship for component *i*, given some DS is given by:

$$\sigma^2_{\ln L_i|DS} = \sigma^2_{mat_i} + \sigma^2_{comptype_i} + \sigma^2_{comp_i} \tag{20}$$

where $\sigma^2_{mat_i}$, $\sigma^2_{comptype_i}$, and $\sigma^2_{comp_i}$ are the uncertainty in the repair cost/duration for material type, component type, and component, respectively, for component *i*. The covariance in the repair cost/duration between two different components, *i* and *j*, is therefore given by:

$$\begin{aligned} Cov\left[\ln L_i|DS, \ln L_j|DS\right] = \delta_{mat_i mat_j}\sigma_{mat_i}\sigma_{mat_j} &+ \delta_{comptype_i comptype_j}\sigma_{comptype_i}\sigma_{comptype_j} \\ &+ \delta_{ij}\sigma_{comp_i}\sigma_{comp_j} \end{aligned} \tag{21}$$

As with the equi-correlated model for the *DS*|*EDP* relationship, the *L*|*DS* equi-correlated model does not consider particular damage states of the components, and only the components themselves. Based on the author's judgement, of the total variance in the damage state uncertainty (Equation (20)), 50% is assumed to be common to specific material types ($\sigma^2_{mat_i}$), 35% common to specific component types ($\sigma^2_{comptype_i}$), and 15% specific to each component ($\sigma^2_{comp_i}$). Based on these assumptions, the correlation coefficient for the *L*|*DS* equi-correlated model is:

$$\rho_{\ln L_i|DS, \ln L_j|DS} = 0.5\delta_{mat_i mat_j} + 0.35\delta_{comptype_i comptype_j} + 0.15\delta_{ij} \tag{22}$$

While the adopted *L*|*DS* equi-correlated model does not consider partial correlations in each of the variables, a clear benefit of using Equation (22), as opposed to the construction correlation data of Touran and Wiser [34] is that different components made of a similar material are not automatically assumed to be perfectly correlated.

CORRELATIONS AND EPISTEMIC UNCERTAINTY

Correlations and epistemic uncertainties are coupled in two ways in a rigorous seismic loss assessment. Firstly, because of the various assumptions made in determining the correlation coefficients, these values have some associated epistemic (knowledge-based) uncertainty. Secondly, epistemic uncertainties in the input values in a seismic loss assessment tend to be correlated at different values of the dependent variables. Neglect of both of these two points can potentially lead to erroneous decision making and a brief discussion is given below concerning these two points.

Epistemic uncertainty in correlation coefficients

There is potentially significant epistemic (knowledge) uncertainty in the correlation coefficients examined to date in this manuscript due to the equi-correlated model assumed and judgement required in determining the magnitude of the variables (in the case of the *DS*|*EDP* and *L*|*DS* relationships) and assumptions made in seismic response modelling (in the case of the $EDP|IM$ relationship). The method by which epistemic uncertainty in correlation coefficients is considered in seismic loss estimation will be partially dependent on how epistemic uncertainty is considered in each of the $EDP|IM$, $DS|EDP$ and $L|DS$ relationships. Here, for brevity, only a single possible approach is considered which is in line with developments to date, and is, the author's opinion, a good compromise between simplicity and accuracy.

Epistemic uncertainties in the $EDP|IM$ relationship are treated in a discrete fashion via the use of logic-trees, much the same as those used for treating epistemic uncertainties in seismic hazard analyses [35]. In this discrete case, the epistemic uncertainty in the correlation coefficients between various EDP's is simply accounted for by computing different correlation coefficients for each set of analyses performed using the different logic tree models. This discrete logic-tree approach is preferred over a continuous treatment of the epistemic uncertainty because of the anticipated complexity in defining the epistemic uncertainties and their correlation with the other EDP epistemic uncertainties in continuous form.

Epistemic uncertainties in both the $DS|EDP$ and $L|DS$ relationships are treated in continuous form by randomly generating the distribution parameters (mean and variance) of the fragility and loss functions [e.g. 30]. Epistemic uncertainty in the values of the correlation coefficients in the $DS|EDP$ and $L|DS$ relationships can be considered by having uncertain variable magnitudes in the generalised equi-correlated model. Hence, for each realization of the magnitude of the variables a correlation coefficient can be computed. This approach is both simple and also ensures that the correlation matrix is strictly semi-positive definite. If the $DS|EDP$ or $L|DS$ correlation matrix is required for the simulation of correlated random numbers in the loss estimation [e.g. 6, 14] then it is possible that the Cholesky decomposition of a non semi-positive definite correlation matrix may contain imaginary numbers.

A uniform probability distribution is assumed for each of the variables in the generalised equi-correlated model. The parameters of the uniform distribution (the lower and upper bounds), were defined as 0.5 and 1.5 times (i.e. ±50% of) the judgement-based value. For example, in the $DS|EDP$ relationship, $\sigma^2_{structure_i}$ is 0.3 x 0.67=0.201 of the total uncertainty, and thus its uniform distribution has lower and upper bounds of 0.101 and 0.302, respectively. Figure 4 illustrates the distribution of the $DS|EDP$ correlation coefficients for two cases based on 10000 Monte Carlo simulations of the random magnitudes of the equi-correlated model variables. It is noted that despite the use of the uniform distribution for the variables: (i) the distribution of the correlation coefficient in Figure 4a is approximately normal (although doubly truncated); (ii) Figure 4b is skewed to the left; and (iii) the variance in Figure 4a is larger than in Figure 4b. All of the above three observations are inline with the Fisher transformation of the correlation coefficient having a normal distribution [36].

Correlations in epistemic uncertainties

In addition to the aleatory uncertainty correlations between different variables which have been discussed in previous sections, there are also correlations between epistemic uncertainties which should be considered in seismic loss estimations.

Bradley [37] discusses the typical correlation structure of epistemic uncertainty in seismic hazard curves. If a logic tree approach [35] is adopted then epistemic uncertainty correlations are implicitly accounted for, while equations are also available if epistemic uncertainties are treated in a parametric form.

Epistemic uncertainties in the $EDP|IM$ relationship resulting from modelling uncertainties in seismic response analysis are starting to gain attention in literature [e.g. 38, 39]. Because of the likely dependence on model configuration it is suggested that correlation between $EDP|IM$ epistemic uncertainties is also handled in a non-parametric logic tree format.

Bradley [30] discussed causes of and methods to determine epistemic uncertainties and their correlations in component fragility functions. As noted by Aslani [3], there is currently no literature suggesting the magnitude of epistemic uncertainty (and their correlations) in component loss functions. Clearly, future work is required in this area, as it is expected the magnitude of the epistemic uncertainty will be significant.

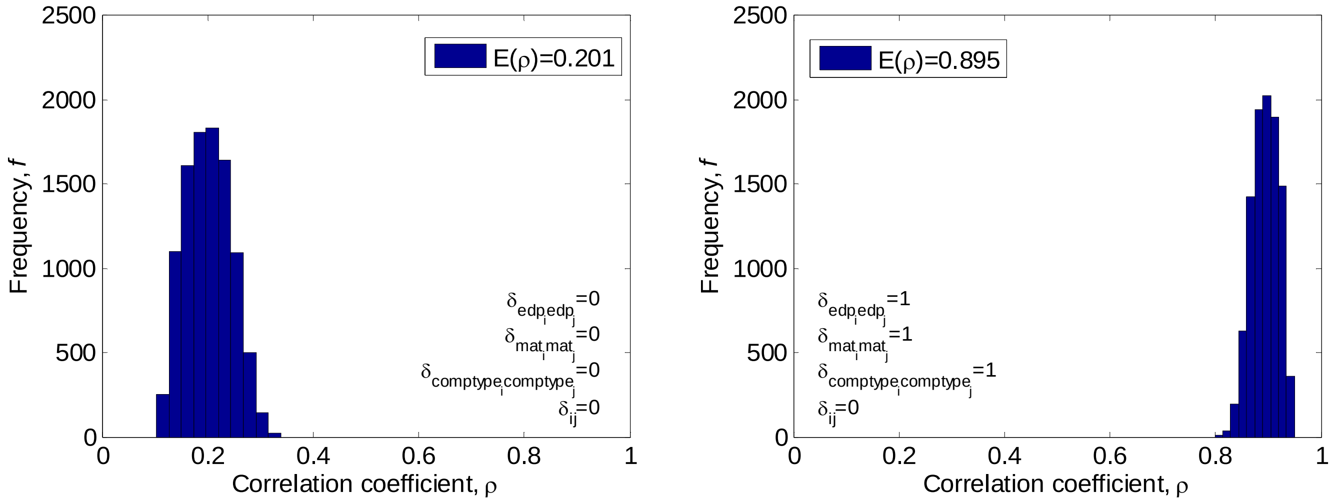

Figure 4: Distribution of the correlation coefficient based on 10000 Monte Carlo simulations with each variable having a uniform distribution with a variation of ±50% of its best-estimate value.

CASE-STUDY SEISMIC LOSS ESTIMATION RESULTS

This section illustrates the effects of five different assumptions regarding the correlation coefficients discussed to date. The seismic loss estimation results presented are for a typical New Zealand 10 storey office building, and include losses resulting from damage of structural, non-structural, and contents components. Details on the seismic hazard, input ground motions and seismic response analyses can be found in Bradley *et al*. [23], while component inventory data is given in Bradley *et al*. [24]. The five different assumptions regarding correlations are: (i) all correlations zero; (ii) all correlations perfect; (iii) all best-estimate partial correlation; (iv) $EDP|IM$ correlations perfect and $DS|EDP$ and $L|DS$ correlations zero; (v) $EDP|IM$ correlations perfect and $DS|EDP$ and $L|DS$ best-estimate partial correlation. Assumptions (i) and (ii) represent the bounding solutions, while assumptions (iv) and (v) represent attempts to approximate the 'correct' assumption (iii) which are less computationally demanding. For simplicity, epistemic uncertainties in the estimated correlation coefficients are not considered in the example to follow.

Total loss given collapse, $L|C$

Figure 5 illustrates the probability density function (pdf) for the total loss given collapse, for three different $L|DS$ correlation assumptions (assumptions (i)-(v) above result in only three unique solutions in this case). As previously noted, in the adopted framework correlation assumptions do not affect the expected collapse loss (Equation (5)) which was $12 million, and the cost distribution is assumed to be lognormal [34]with variance obtained from Equation (6). Figure 5 illustrates that the lognormal standard deviation assuming perfect correlations is three times larger than assuming no correlations. Using partial correlations yields a standard deviation which is 24% larger than that assuming no correlation assumption, similar to that given by Aslani [3].

Comparison of the magnitude of the lognormal standard deviations for the total loss given collapse indicates that the results presented here are somewhat smaller than those given by Aslani [3] (which were lognormal standard deviations of 0.24, 0.4, and 0.65 for none, partial, and perfect correlations, respectively). This is largely due to significantly larger number of components and component types considered compared with Aslani [3], and as Aslani [3] illustrates, the lognormal standard deviation will reduce as the number of partially correlated components increases. This same logic applies to the lognormal standard deviations for the $L|IM,NC$ and $L|IM$ relationships discussed below.

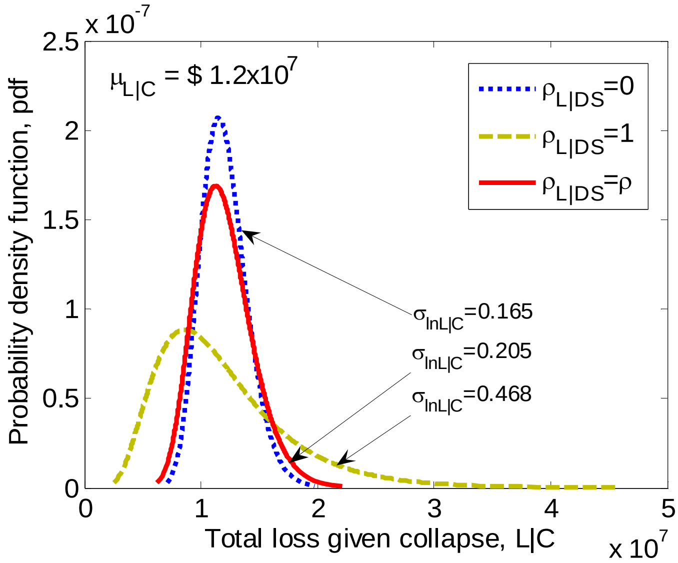

Figure 5: Distribution of total loss given structural collapse with various correlation assumptions

Total loss given no collapse, *L|IM,NC* and total loss, *L|IM*

Figure 6a and 6b illustrate the lognormal standard deviation in the total loss given no collapse and the total loss, respectively, while Figure 6c and 6d give the corresponding error ratios compared to the case of all correlations allowed to be partial. Firstly, it is noted from Figure 6a and 6c that the bounding assumptions of none and perfect correlations are typically in error by over 50%, and the error is relatively constant as a function of the ground motion IM (which is peak ground velocity, *PGV*). Secondly, both approximations (iv) and (v) result in standard deviations which are in close agreement with the best-estimate solution.

The standard deviation values in Figure 6b are similar those in Figure 6a for small *PGV*, but tend to the standard deviation given collapse (Figure 5), for increasing *PGV* values as indicated by Equation (2). The local maxima in the lognormal standard deviation of the total loss, $\sigma_{\ln L|IM}$, for the case of no correlations occurs due to the contribution of the third term on the right hand side of Equation (2). This term is most significant at the *PGV* value for which $(1-P_{C|IM})P_{C|IM}$ is maximised (i.e. $P_{C|IM}=0.5$). The effect of this third term for the other four correlation assumptions is less pronounced because in these cases $\sigma_{\ln L|C}$ is significantly less than $\sigma_{\ln L|IM,NC}$. Figure 6d illustrates that, for this particular structure, the upper bound assumption of perfect correlations over-approximates $\sigma_{\ln L|IM}$ by about 50% over a wide range of *PGV*, while the error associated with the assumption of no correlations is also up to 50% for small *PGV* values, and decreases with increasing *PGV*.

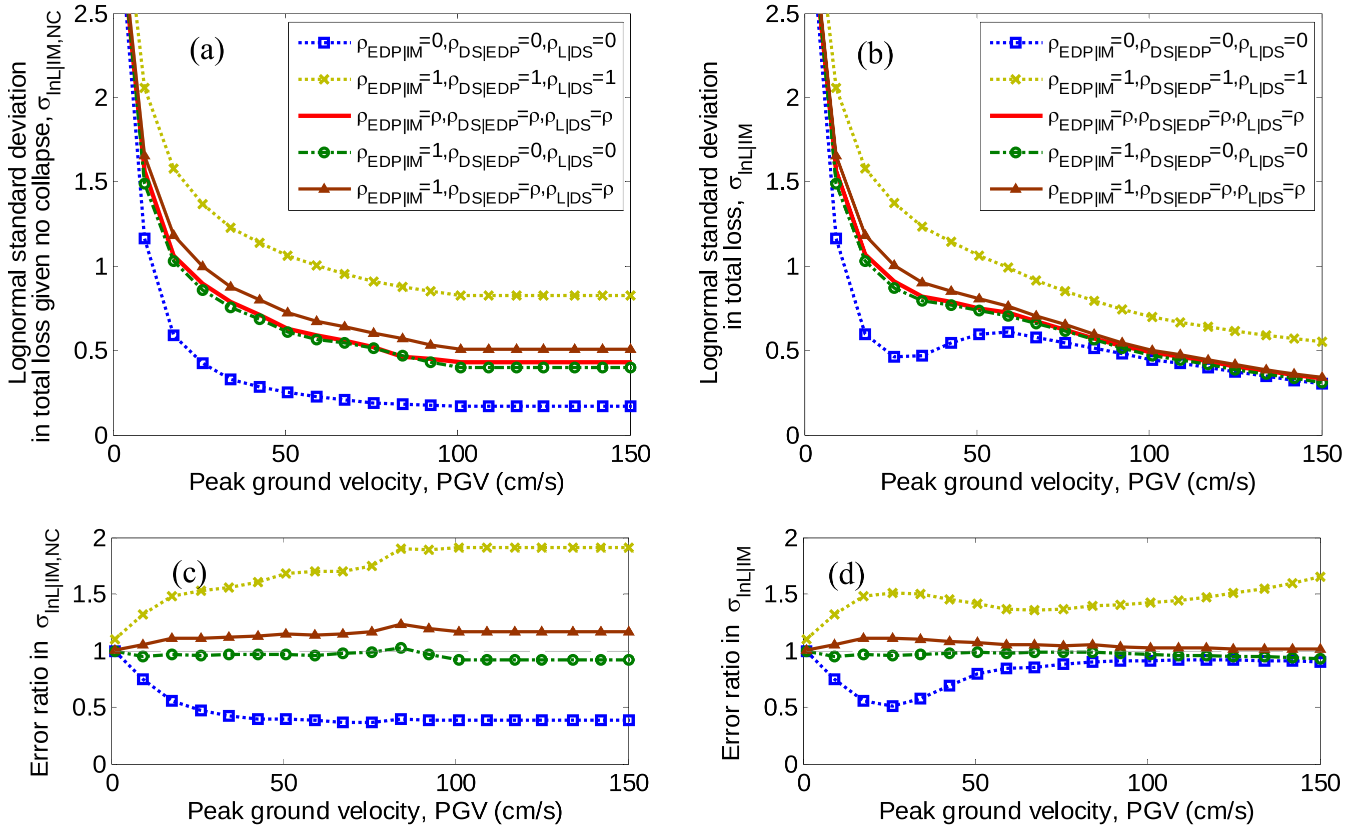

Figure 6: Effect of correlation assumptions on the dispersion in the total loss: (a) lognormal standard deviation in the total loss given no collapse; (b) lognormal standard deviation in total loss; (c) error ratios in the dispersion in the total loss given no collapse; and (d) error ratios in the dispersion in the total loss

sLoss hazard, P_L.

Figure 7a illustrates the loss hazard curve for the case study structure based on the five different correlation approximations, while Figures 7b and 7c illustrate the error ratios in loss and annual probability of exceedance, respectively. Mathematical details of the loss hazard computation can be found in References [3, 16]. Figure 7b illustrates that for exceedance probabilities greater than 0.01 the bounding solutions of zero and perfect correlations yield errors in the loss of around 50%. For this same P_L range, which correspond to losses less than $1\text{x}10^6$ (i.e. \$1 million), the error for the zero correlation assumption is larger than 20%, while for the perfect correlation assumption the error is up to 20%. For losses in excess of $2\text{x}10^6$ (with corresponding probabilities of exceedance less than $2\text{x}10^{-3}$), the trends are reversed with the zero correlation assumption giving an under-prediction and the perfect correlation assumption giving an over-prediction. For losses larger than the mean value of the collapse loss (i.e. $1.2\text{x}10^7$) the loss hazard curve is particularly sensitive to the assumption of the $L|DS$ correlation with error ratios greater than 2 when perfect $L|DS$ correlations were assumed and less than 0.5 when zero $L|DS$ correlations were assumed.

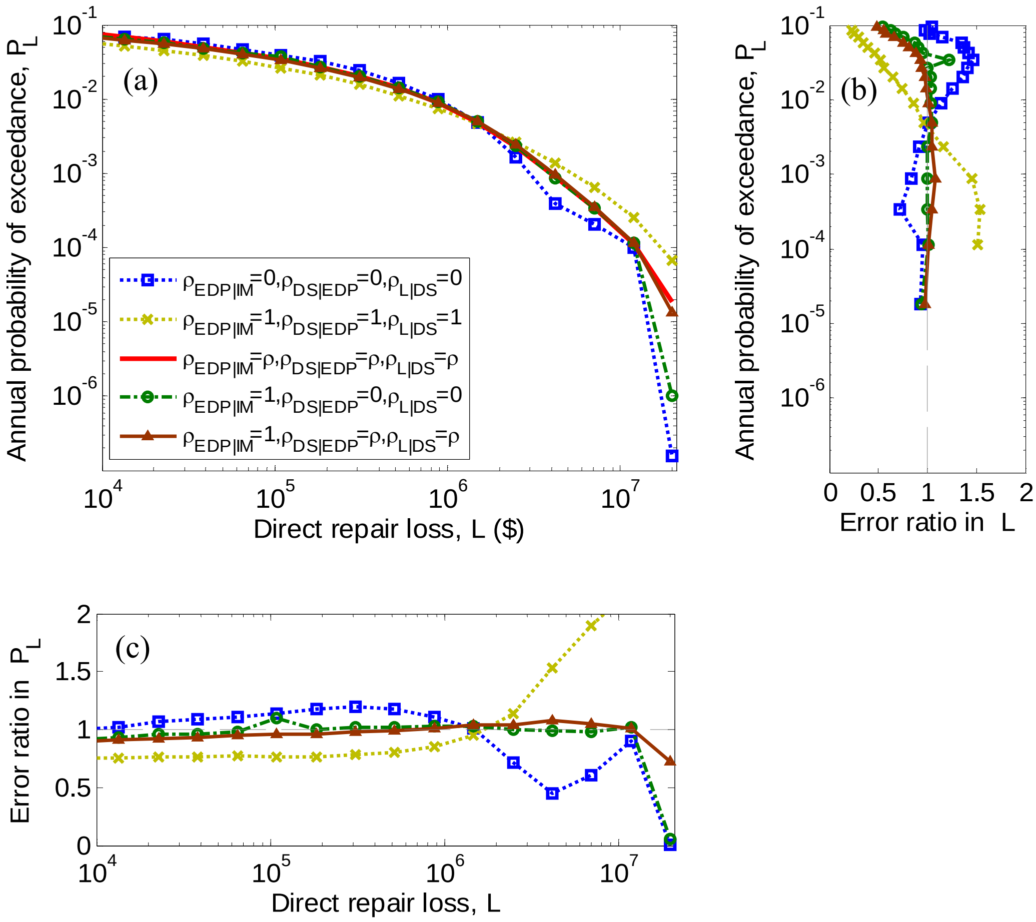

Figure 7: (a) Resulting loss hazard curves based on different assumptions regarding correlations; (b) error ratio in loss, L; and (c) error ratio in annual probability of exceedance, P_L.

Computational demand

Table 1 presents the computational times required when in performing the seismic loss assessment on a Pentium 4 processor with 3.0 GHz CPU and 512 MB RAM using the seismic loss assessment tool (SLAT) [40], which utilizes the magnitude-oriented adaptive quadrature algorithm [41]. As discussed by Bradley and Lee [10], and evident in Table 1, the effect of non-zero correlations drastically increases the computational demand to perform the analysis. Table 1 illustrates that the loss hazard involves approximately 20 times more computational time than the $L|IM$ relationships, and it can be seen that the using perfect and partial correlations requires approximately 50-times and 3200-times more computational time, respectively, than the assumption of no-correlations. The fourth and fifth rows of Table 1 which correspond to correlation approximations (iv) and (v), respectively, illustrate that by assuming perfect $EDP|IM$ correlations and either zero or partial correlations for the $DS|EDP$ and $L|DS$ relationships, the computational time to run the analysis is significantly reduced compared to assuming all correlations are partial. The small difference in computational times for correlation assumptions (iv) and (v) illustrate that the additional computational demand due to the $DS|EDP$ and $L|DS$ correlations is minimal compared to considering $EDP|IM$ correlations. By assuming perfect $EDP|IM$ correlations the double-integral in Equation (9) reduces to a single integral as discussed by Bradley and Lee [10], which significantly reduces the computational time. Since correlation approximation (v) uses perfect $EDP|IM$ correlations and partial $DS|EDP$ and $L|DS$ correlations then it will always give a larger value for the lognormal standard deviation in the total loss (i.e. both $\sigma_{\ln L|IM,NC}$ and $\sigma_{\ln L|IM}$), than using all partial correlations. Thus, based on the accuracy of correlation assumption (iv), it may be viewed as a good approximation if it is deemed that the computational demand associated with assuming all partial correlations is excessive.

Table 1: Computational times for seismic loss analyses

Correlation assumption*		Loss vs. IM	Loss vs. P_L
(i)	$\rho_{EDP\|IM} = 0, \rho_{DS\|EDP} = 0, \rho_{L\|DS} = 0$	2.6 sec	52 sec
(ii)	$\rho_{EDP\|IM} = 1, \rho_{DS\|EDP} = 1, \rho_{L\|DS} = 1$	130 sec (2.1 min)	2500 sec (41 min)
(iii)	$\rho_{EDP\|IM} = \rho, \rho_{DS\|EDP} = \rho, \rho_{L\|DS} = \rho$	8200 sec (2.3 hr)	180000 sec (49 hr)
(iv)	$\rho_{EDP\|IM} = 1, \rho_{DS\|EDP} = 0, \rho_{L\|DS} = 0$	130 sec (2.1 min)	2800 sec (46 min)
(v)	$\rho_{EDP\|IM} = 1, \rho_{DS\|EDP} = \rho, \rho_{L\|DS} = \rho$	190 sec (3.2 min)	3700 sec (1.0 hr)

*Correlation coefficient values of 0, 1, and ρ correspond to zero, perfect, and partial correlations, respectively.

CONCLUSIONS

The consideration of component correlations in seismic loss estimation has been limited by methodological tractability, increased computational demand, and a paucity of data for their computation. This paper has presented a tractable and computationally efficient seismic loss estimation methodology in which correlations can be considered. Methods to determine the necessary correlations were discussed, particularly those which can be used in the absence of sufficient empirical data and rely somewhat on judgement. The effects of various assumptions regarding correlations were illustrated via application to a case-study office structure. It was observed that certain correlation assumptions can lead to errors in excess of 50% in the lognormal standard deviation in the loss given intensity and loss hazard relationships, while full consideration of partial correlations requires in excess of 50-times more computational time than other correlation assumptions.

ACKNOWLEDGEMENTS

Financial support of the first author from the New Zealand Tertiary Education Commission Bright Futures scheme is appreciated.

REFERENCES:

[1] Goulet CA, Haselton CB, Mitrani-Reiser J, Beck JL, Deierlein GG, Porter K, and Stewart JP. Evaluation of the seismic performance of a code-conforming reinforced-concrete frame building - from seismic hazard to collapse safety and economic losses. *Earthquake Engineering and Structural Dynamics* 2007; **36**(13): 1973-1997, DOI: 10.1002/eqe.694.

[2] Mitrani-Reiser J. An Ounce of Prevention: Probabilistic Loss Estimation for Performance-based Earthquake Engineering. Ph.D. Thesis, California Institute of technology, 2007, 173.

[3] Aslani H. Probabilistic earthquake loss estimation and loss disaggregation in buildings. Ph.D. Thesis, John A. Blume Earthquake Engineering Centre, Dept. of Civil and Environmental Engineering Stanford University, 2005, 382.
[4] Porter KA and Kiremidjian AS. Assembly-based vulnerability and its uses in seismic performance evaluation and risk-management decision-making. John A.Blume Earthquake Engineering Center, Stanford, CA, 2001. 216.
[5] Iervolino I, Giorgio M, and Manfredi G. Expected loss-based alarm threshold set for earthquake early warning systems. *Earthquake Engineering and Structural Dynamics* 2007; **36**(9): 1151-1168, DOI: 10.1002/eqe.675.
[6] ATC-58. Guidelines for seismic performance assessment of buildings: ATC-58 35% Draft. Applied Technology Council, 2007. http://www.atcouncil.org/atc-58.shtml (last accessed 19/11/08)
[7] Lee R and Kiremidjian AS. Uncertainty and correlation for loss assessment of spatially distributed systems. *Earthquake Spectra* 2007; **23**(4): 753-770, DOI: 10.1193/1.2791001.
[8] Baker JW. Introducing correlation among fragility functions for multiple components, in *14th World Conference on Earthquake Engineering*, Beijing, China, 2008, 8.
[9] Baker JW and Cornell CA. Uncertainty propagation in probabilistic seismic loss estimation. *Structural Safety* 2008; **30**(3): 236-252.
[10] Bradley BA and Lee DS. Accuracy of approximate methods of uncertainty propagation in loss estimation. *Structural Safety* 2009: (in press), DOI: 10.1016/j.strusafe.2009.04.001.
[11] Cornell CA and Krawinkler H. Progress and challenges in seismic performance assessment. *PEER Center News* 2000; **3**(2).
[12] Ang AHS and Tang WH. *Probability Concepts in Engineering Planning and Design* vol. Volume I – Basic Principles. John Wiley & Sons, Inc., 1975; 406.
[13] Kircher CA, Reitherman RK, Whitman RV, and Arnold C. Estimation of earthquake losses to buildings. *Earthquake Spectra* 1997; **13**(4): 703-720.
[14] Porter KA, Kiremidjian AS, and LeGrue JS. Assembly-Based Vulnerability for Buildings and Its Use in Performance Evaluation. *Earthquake Spectra* 2001; **17**(2): 291-312.
[15] Porter K, Kennedy R, and Bachman R. Creating Fragility Functions for Performance-based Earthquake Engineering. *Earthquake Spectra* 2007; **23**(2): 471-489.
[16] Bradley BA, Dhakal RP, Cubrinovski M, MacRae GA, and Lee DS. Seismic loss estimation for efficient decision making. *Bulletin of the New Zealand Society for Earthquake Engineering* 2008.
[17] Bazzurro P. Probabilistic seismic demand analysis. Stanford University, Stanford, CA, 1998. 329. http://www.stanford.edu/group/rms/
[18] Aslani H and Miranda E. Probability-based Seismic Response Analysis. *Engineering Structures* 2005; **27**(8): 1151-1163.
[19] Aslani H and Miranda E. Fragility assessment of slab-column connections in existing non-ductile reinforced concrete buildings. *Journal of Earthquake Engineering* 2005; **9**(6): 777–804.
[20] Genz A. Numerical computation of rectangular bivariate and trivariate normal and t probabilities. *Statistics and Computing* 2004; **14**(3): 251-260.
[21] Drezner Z and Wesolowsky GO. On the computation of the bivariate normal inegral. *Journal of Statistical and Computational Simulation* 1990; **35**: 101-107.
[22] Chopra AK and Goel R. A modal pushover analysis procedure to estimate seismic demands for unsymmetric-plan buildings. *Earthquake Engineering and Structural Dynamics* 2004; **33**: 903–927.
[23] Bradley BA, Dhakal RP, Cubrinovski M, and MacRae GA. Prediction of spatially distributed seismic demands in structures: ground motion and structural response. *Earthquake Engineering and Structural Dynamics* 2008: (submitted).

[24] Bradley BA, Dhakal RP, Cubrinovski M, and MacRae GA. Prediction of spatially distributed seismic demands in structures: from structural response to loss estimation. *Earthquake Engineering and Structural Dynamics* 2008: (submitted).
[25] Fajfar P. Structural analysis in earthquake engineering – A breakthrough of simplified nonlinear methods, in *12th European Conference on Earthquake Engineering*, London, UK., 2002.
[26] Baker JW and Cornell CA. Uncertainty Specification and Propagation for Loss Estimation using FOSM Method. Peer Report No 2003/07, 2003. 100. http://peer.berkeley.edu/publications/peer_reports.html
[27] Iervolino I, Manfredi G, and Cosenza E. Ground motion duration effects on nonlinear seismic response. *Earthquake Engineering and Structural Dynamics* 2006; **35**(1): 21-38, DOI: 10.1002/eqe.529.
[28] Luco N and Cornell CA. Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. *Earthquake Spectra* 2007; **23**(2): 357-392, DOI: 10.1193/1.2723158.
[29] Ditlevsen O. *Uncertainty Modeling: with Applications to Multidimensional Civil Engineering Systems*. McGraw-Hill International Book Co.: New York; London, 1981; 412.
[30] Bradley BA. Epistemic uncertainty in component fragility functions. *Earthquake Spectra* 2009: (in press).
[31] Porter K. Fragility of Hydraulic Elevators for Use in Performance-Based Earthquake Engineering. *Earthquake Spectra* 2007; **23**(2): 459-469.
[32] Touran A. Probabilistic cost estimating with subjective correlations. *Journal of Construction Engineering and Management* 1993; **119**(1): 58-71.
[33] Hudak D and Maxwell M. A macro approach to estimating correlated random variables in engineering production projects. *Construction Management and Economics* 2007; **25**: 883-892.
[34] Touran A and Wiser EP. Monte carlo technique with correlated random variables. *Journal of Construction Engineering and Management* 1992; **118**(2): 258-272.
[35] Kulkarni RB, Youngs RR, and Coppersmith KJ. Assessment of confidence intervals for results of seismic hazard analysis. *8th World Conference on Earthquake Engineering,*, San Francisco, CA, 1984.
[36] Kutner M, Nachtsheim C, Neter J, and Li W. *Applied Linear Statistical Models*. Fifth Edition ed. McGraw-Hill/Irwin: New York, 2005; 1396.
[37] Bradley BA. Seismic hazard epistemic uncertainty in the San Francisco bay area and its role in performance-based assessment. *Earthquake Spectra* 2009: (in press).
[38] Dolsek M. Incremental dynamic analysis with consideration of modelling uncertainties. *Earthquake Engineering and Structural Dynamics* 2009, DOI: 10.1002/eqe.869.
[39] Kramer SL, Arduino P, and Shin H. Using OpenSees for performance-based evaluation of bridges on liquefiable soils. Pacific Earthquake Engineering Research Centre, 2008. 198.
[40] Bradley BA. User manual for SLAT: Seismic Loss Assessment Tool version 1.14. *University of Canterbury Research Report No.2009-01*,Department of Civil and Natural Resources Engineering, University of Canterbury, Christchurch, New Zealand, 2009. 94.
[41] Bradley BA, Lee DS, Broughton R, and Price C. Efficient Evaluation of Performance-based Earthquake Engineering Equations. *Structural Safety* 2009; **31**(1): 65-74.